

Stat 521
11-25-25

Defn: Let $f(x)$ be a density function
with parameters $\vec{\theta} = (\theta_1, \theta_2, \dots, \theta_d)$.

(1)

Then $f(x|\vec{\theta})$ [or $f_{\vec{\theta}}(x)$] is an
exponential family of densities if you can
write $f(x|\vec{\theta}) = h(x)c(\vec{\theta})e^{\sum_{i=1}^d w_i(\vec{\theta})t_i(x)}$,
 $-\infty < x < \infty$.

If $d=k$, then the family is full.

If $d < k$, then the family is curved.

(2)

Example: $\text{Binomial}(n, p)$ with n known

$$f(x|p) \text{ [or } f_p(x)] = \binom{n}{x} p^x (1-p)^{n-x},$$

$$x = 0, 1, \dots, n$$

$$\text{Let } I(x) = \begin{cases} 1 & \text{if } x = 0, 1, \dots, n \\ 0 & \text{otherwise} \end{cases}$$

$$f(x|p) = \binom{n}{x} I(x) (1-p)^n e^{[x \ln p - x \ln(1-p)]}$$

$$-\infty < x < \infty$$

$$h(x) = \binom{n}{x} I(x) \quad w_1(\vec{\theta}) = \ln p \quad t_1(x) = x$$

$$c(\vec{\theta}) = (1-p)^n \quad w_2(\vec{\theta}) = \ln(1-p) \quad t_2(x) = -x$$

$$f(x|p) = \binom{n}{x} I(x) (1-p)^n e^{[x \cdot \ln(\frac{p}{1-p})]} \quad (3)$$

$$h(x) = \binom{n}{x} I(x)$$

$$c(\vec{\theta}) = (1-p)^n$$

$$w_1(\vec{\theta}) = \ln\left(\frac{p}{1-p}\right)$$

$$t_1(x) = x$$

$$-\infty < x < \infty$$

$\therefore \text{Bin}(n, p)$ with n known is a full exponential family.

Example: $N(\mu, \sigma^2)$

$$f(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$$

$$f(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x^2 - 2\mu x + \mu^2)} \quad (4)$$

$$-\infty < x < \infty$$

$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} e^{\frac{\mu x}{\sigma^2}} e^{-\frac{\mu^2}{2\sigma^2}}$$

$$h(x) = \frac{1}{\sqrt{2\pi}} \quad c(\vec{\theta}) = \frac{1}{\sigma} e^{-\frac{\mu^2}{2\sigma^2}}$$

$$w_1(\vec{\theta}) = -\frac{1}{2\sigma^2} \quad t_1(x) = x^2$$

$$w_2(\vec{\theta}) = \frac{\mu}{\sigma^2} \quad t_2(x) = x$$

$\therefore N(\mu, \sigma^2)$ is a full exponential family

(5)

Example: Unit (a,b)

$$f(x|a,b) = \frac{1}{b-a} \quad a < x < b$$

$$= \frac{1}{b-a} I(x) \quad -\infty < x < \infty$$

$$\text{where } I(x) = \begin{cases} 1 & a < x < b \\ 0 & \text{otherwise} \end{cases}$$

This is not an exponential family.

(6)

Theorem: Let $f(x)$ be a pdf.

Let μ be any real constant.

Let σ be any positive real constant.

$$\text{Then } \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right) = g(x|\mu, \sigma)$$

is also a valid pdf.

Proof: $g(x|\mu, \sigma) \geq 0 \quad \checkmark$

$$\int_{-\infty}^{\infty} g(x) dx = \int_{-\infty}^{\infty} \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right) dx$$

$$\text{Let } u = \frac{x-\mu}{\sigma} \\ du = \frac{1}{\sigma} dx$$

$$= \int_{-\infty}^{\infty} f(u) du = 1 \quad \checkmark$$

(7)

Defn. Let $f(x)$ be a pdf.

Let $g(x|\mu) = f(x-\mu)$. Then $g(x|\mu)$ is a location family of densities. And μ is a location parameter.

Let $g(x|\sigma) = \frac{1}{\sigma} f(\frac{x}{\sigma})$, for $\sigma > 0$.

Then $g(x|\sigma)$ is a scale family.

σ is a scale parameter.

(8)

Let $g(x|\mu, \sigma) = \frac{1}{\sigma} f(\frac{x-\mu}{\sigma})$ for $\sigma > 0$.

Then $g(x|\mu, \sigma)$ is a location-scale family.

Example: $N(\mu, \sigma^2)$

$$g(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$\therefore N(\mu, \sigma^2)$ is a location-scale family

$$= \frac{1}{\sigma} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

where $z = \frac{x-\mu}{\sigma}$

$$= \frac{1}{\sigma} f\left(\frac{x-\mu}{\sigma}\right) \text{ where } f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

(9)

Example: Gamma(α, β)

$$q(x | \alpha, \beta) = \frac{1}{\Gamma(\alpha) \beta^\alpha} x^{\alpha-1} e^{-x/\beta}$$

$$= \frac{1}{\beta} \frac{1}{\Gamma(\alpha)} \left(\frac{x}{\beta}\right)^{\alpha-1} e^{-x/\beta}$$

$$= \frac{1}{\beta} f\left(\frac{x}{\beta}\right) \quad \text{where} \quad f(z) = \frac{1}{\Gamma(\alpha)} z^{\alpha-1} e^{-z}$$

$\therefore \beta$ is a scale parameter

(10)

Example: Cauchy(θ)

$$q(x | \theta) = \frac{1}{\pi} \frac{1}{1 + (x - \theta)^2} \quad -\infty < x < \infty$$

$$= \frac{1}{\pi} \frac{1}{1 + z^2} \quad \text{where} \quad z = x - \theta$$

$$= f(x - \theta) \quad \text{where} \quad f(z) = \frac{1}{\pi} \frac{1}{1 + z^2}$$

$\therefore \theta$ is a location parameter

Random vectors

(11)

A random vector $\vec{X}$ is a function from the sample space S into $\mathbb{R}^n$.

(i.e. it is a vector whose components are random variables)

A random vector is discrete if its range in $\mathbb{R}^n$ is finite or countably infinite.

For $A \subset \mathbb{R}^n$,

(12)

$$P(A) = P[\vec{X} \in A]$$

Defn: For a discrete random vector, the joint probability mass function is

$$p(x_1, x_2, \dots, x_n) = P[X_1 = x_1 \cap X_2 = x_2 \cap \dots \cap X_n = x_n]$$

$$\text{So } P(A) = \underbrace{\sum \sum \dots \sum}_{(x_1, \dots, x_n) \in A} p(x_1, x_2, \dots, x_n)$$

(13)

Example: 4 slips of paper, numbered 1, 2, 3, 4

Draw 2 w.o.r.

Let $X_1 = \min$ $X_2 = \max$

$$S = \{12, 13, 14, 21, 23, 24, 31, 32, 34, 41, 42, 43\}$$

$p(x_1, x_2)$		X_2			
		2	3	4	
X_1	1	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{3}{6}$
	2	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$
	3	0	0	$\frac{1}{6}$	$\frac{1}{6}$
		$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	1

Theorem:
$$p_1(x_1) = \sum_{\text{all } x_2} p(x_1, x_2)$$

(14)

Proof:
$$p_1(x_1) = P[X_1 = x_1]$$

$$= P[X_1 = x_1 \cap -\infty < X_2 < \infty]$$

$$= \sum_{\text{all } x_2} P[X_1 = x_1 \cap X_2 = x_2]$$

$$= \sum_{\text{all } x_2} p(x_1, x_2)$$